

Rare loss-of-function variants in type I IFN immunity genes are not associated with severe COVID-19

Gundula Povysil,¹ Guillaume Butler-Laporte,^{2,3} Ning Shang,⁴ Chen Wang,⁴ Atlas Khan,⁴ Manal Alaamery,^{5,6} Tomoko Nakanishi,^{2,7,8} Sirui Zhou,² Vincenzo Forgetta,² Robert J.M. Eveleigh,^{9,10} Mathieu Bourgey,^{9,10} Naveed Aziz,¹¹ Steven J.M. Jones,¹¹ Bartha Knoppers,¹¹ Stephen W. Scherer,¹¹ Lisa J. Strug,¹¹ Pierre Lepage,¹² Jiannis Ragoussis,^{7,12} Guillaume Bourque,^{7,10,12} Jihad Alghamdi,¹³ Nora Aljawini,¹⁴ Nour Albes,¹⁴ Hani M. Al-Afghani,¹⁵ Bader Alghamdi,⁵ Mansour S. Almutairi,⁵ Ebrahim Sabri Mahmoud,¹⁶ Leen Abu-Safieh,¹⁷ Hadeel El Bardisy,¹⁷ Fawz S. Al Harthi,¹⁷ Abdulraheem Alshareef,¹⁸ Bandar Ali Suliman,¹⁸ Saleh A. Alqahtani,^{19,20} Abdulaziz Almalik,²¹ May M. Alrashed,²² Salam Massadeh,^{5,6} Vincent Mooser,⁷ Mark Lathrop,^{7,11,12} Mohamed Fawzy,¹⁷ Yaseen M. Arabi,¹⁶ Hamdi Mbarek,²³ Chadi Saad,²³ Wadha Al-Muftah,²³ Junghyun Jung,^{24,25} Serghei Mangul,²⁴ Radja Badji,²³ Asma Al Thani,²³ Said I. Ismail,²³ Ali G. Gharavi,^{1,4,26} Malak S. Abedalthagafi,¹⁷ J. Brent Richards,^{2,3,8,27} David B. Goldstein,^{1,28} and Krzysztof Kiryluk^{1,4}

¹Institute for Genomic Medicine, Columbia University, New York, New York, USA. ²Lady Davis Institute for Medical Research, Montréal, Québec, Canada. ³Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montréal, Québec, Canada. ⁴Division of Nephrology, Department of Medicine, Vagelos College of Physicians & Surgeons, Columbia University, New York, New York, USA. ⁵Developmental Medicine Department, King Abdullah International Medical Research Center, King Saud Bin Abdulaziz University for Health Sciences, King Abdulaziz Medical City, Ministry of National Guard Health Affairs, Riyadh, Saudi Arabia. ⁶Saudi Human Genome Project at King Abdulaziz City for Science and Technology, Riyadh, Saudi Arabia. ⁷Department of Human Genetics, McGill University, Montréal, Québec, Canada. ⁸Kyoto-McGill International Collaborative School in Genomic Medicine, Graduate School of Medicine, Kyoto University, Kyoto, Japan. ⁹Canadian Centre for Computational Genomics, McGill University, Montréal, Québec, Canada. ¹⁰McGill Genome Center, McGill University, Montréal, Québec, Canada. ¹¹Canadian COVID Genomics Network, HostSeq Project, Canada. ¹²Canadian Centre for Computational Genomics, McGill University, Montréal, Québec, Canada. ¹³Saudi Biobank, King Abdullah International Medical Research Center, King Saud bin Abdulaziz University for Health Sciences, Ministry of National Guard Health Affairs, Riyadh, Saudi Arabia. ¹⁴KACST-BWH Centre of Excellence for Biomedicine, Joint Centers of Excellence Program, King Abdulaziz City for Science and Technology, Riyadh, Saudi Arabia. ¹⁵Laboratory Department, Security Forces Hospital, General Directorate of Medical Services, Ministry of Interior, Makkah, Saudi Arabia. ¹⁶Ministry of the National Guard Health Affairs, King Abdullah International Medical Research Center and King Saud Bin Abdulaziz University for Health Sciences, Riyadh, Saudi Arabia. ¹⁷Genomics Research Department, Saudi Human Genome Project, King Fahad Medical City and King Abdulaziz City for Science and Technology, Riyadh, Saudi Arabia. ¹⁸College of Applied Medical Sciences, Taibah University, Madina, Saudi Arabia. ¹⁹Liver Transplant Unit, King Faisal Specialist Hospital and Research Centre, Riyadh, Saudi Arabia. ²⁰Division of Gastroenterology and Hepatology, Johns Hopkins University, Baltimore, Maryland, USA. ²¹Life Science and Environmental Institute, King Abdulaziz City for Science and Technology, Riyadh, Saudi Arabia. ²²Department of Clinical Laboratory Sciences, College of Applied Medical Sciences, King Saud University, Riyadh, Saudi Arabia. ²³Qatar Genome Program, Qatar Foundation Research, Development and Innovation, Qatar Foundation, Doha, Qatar. ²⁴Department of Clinical Pharmacy, School of Pharmacy, University of Southern California, Los Angeles, California, USA. ²⁵Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California, USA. ²⁶Center for Precision Medicine and Genomics, Department of Medicine, Vagelos College of Physicians & Surgeons, Columbia University, New York, New York, USA. ²⁷Department of Twin Research, King's College London, London, United Kingdom. ²⁸Department of Genetics & Development, Columbia University, New York, New York, USA.

A recent report found that rare predicted loss-of-function (pLOF) variants across 13 candidate genes in TLR3- and IRF7-dependent type I IFN pathways explain up to 3.5% of severe COVID-19 cases. We performed whole-exome or whole-genome sequencing of 1,864 COVID-19 cases (713 with severe and 1,151 with mild disease) and 15,033 ancestry-matched population controls across 4 independent COVID-19 biobanks. We tested whether rare pLOF variants in these 13 genes were associated with severe COVID-19. We identified only 1 rare pLOF mutation across these genes among 713 cases with severe COVID-19 and observed no enrichment of pLOFs in severe cases compared to population controls or mild COVID-19 cases. We found no evidence of association of rare LOF variants in the 13 candidate genes with severe COVID-19 outcomes.

Introduction

Host genetic factors related to COVID-19 susceptibility and outcomes are of great interest, as their identification could elucidate the mechanisms of SARS-CoV-2 infection and severity, there-

► **Related Commentary:** <https://doi.org/10.1172/JCI149459>

Conflict of interest: The authors have declared that no conflict of interest exists.

Copyright: © 2021, American Society for Clinical Investigation.

Submitted: January 19, 2021; **Accepted:** May 25, 2021; **Published:** July 15, 2021.

Reference information: *J Clin Invest.* 2021;131(14):e147834.

<https://doi.org/10.1172/JCI147834>.

by providing clues about potential therapeutic targets. Recently, Zhang et al. (1) reported deleterious mutations in 13 candidate genes involved in type I IFN immunity to be associated with COVID-19 severity. The authors reported a significant enrichment in predicted loss-of-function (pLOF) variants in 659 severe COVID-19 cases relative to 534 controls with asymptomatic or benign infection using a burden test under a dominant model of inheritance (odds ratio [OR] = 8.28 [95% CI 1.04–65.64], $P = 0.01$). In addition to the 9 pLOF variants in cases used in the burden test, they detected 109 missense or in-frame indels and tested them experimentally in overexpression systems; 24 variants (including all pLOFs) were found to be hypomorphic and were detected in 23 patients of both

Table 1. Study cohorts

	COVID-19 cases	Severe COVID-19	Mild COVID-19	Population controls
Columbia	1,153	480	673	9,589
Quebec	220	62	158	313
Saudi Arabia	237	148	89	0
Qatar	254	23	231	5,131
Total	1,864	713	1,151	15,033

sexes and various ages (17–77 years). The authors reported the presence of 1 pLOF variant in controls but did not provide the total number of other functional variants, or whether variants identified in controls were experimentally tested. They concluded that rare deleterious variants in the TLR3-, IRF7-, and IRF9-dependent type I IFN pathway genes explain up to 3.5% of severe COVID-19 cases.

We aimed to replicate these findings using independent data sets. We note that an exact replication of all the reported associa-

tions is not possible because the study design of Zhang et al. (1) does not permit a formal statistical comparison of the same variant types in cases and controls. Missense variants were reported as functionally characterized in cases but not controls, so the counts of functionally compromised missense variants cannot be evaluated in a statistical test. We therefore performed association tests on all rare variation that is apparently functional, and specifically focused on pLOF variants, applying identical rules to case and control variants, as is

appropriate in such analyses.

Results and Discussion

Our first cohort was recruited by the Columbia University COVID-19 Biobank (Tables 1 and 2). We performed exome sequencing of 1,153 cases of diverse ancestries; 480 of 1,153 cases had severe COVID-19 that led to death or required intubation due to respiratory failure (Table 2 and Supplemental Methods;

Table 2. Cohort characteristics

	Columbia University COVID-19 Biobank		Biobanque Québec COVID-19		Saudi Arabia COVID-19 Biobank		Qatar Genome Program COVID-19	
	Severe n = 480	Mild n = 673	Severe n = 62	Mild n = 158	Severe n = 148	Mild n = 89	Severe n = 23	Mild n = 231
Age								
0–9	3 (0.62%)	9 (1.34%)	0	0	0	0	0	0
10–19	9 (1.88%)	5 (0.74%)	0	0	1 (0.68%)	2 (2.24%)	0	2 (0.87%)
20–29	5 (1.04%)	34 (5.05%)	2 (3.23%)	6 (3.80%)	0	15 (16.85%)	1 (4.35%)	57 (24.68%)
30–39	16 (3.33%)	66 (9.81%)	2 (3.23%)	10 (6.33%)	7 (4.73%)	36 (40.45%)	3 (13.04%)	75 (32.47%)
40–49	28 (5.83%)	60 (8.92%)	2 (3.23%)	13 (8.23%)	10 (6.76%)	23 (25.84%)	7 (30.43%)	45 (19.48%)
50–59	62 (12.92%)	124 (18.42%)	9 (14.52%)	27 (17.09%)	18 (12.16%)	4 (4.49%)	7 (30.43%)	37 (16.02%)
60–69	97 (20.21%)	137 (20.36%)	13 (20.97%)	20 (12.66%)	48 (32.43%)	8 (8.98%)	5 (21.74%)	11 (4.76%)
70–79	133 (27.71%)	130 (19.32%)	17 (27.42%)	24 (15.19%)	47 (31.76%)	1 (1.12%)	0	4 (1.73%)
80–89	89 (18.54%)	72 (10.70%)	11 (17.74%)	40 (25.32%)	17 (11.49%)	0	0	0
≥90	31 (6.46%)	23 (3.42%)	5 (8.06%)	15 (9.49%)	1 (0.68%)	0	0	0
Unknown	7 (1.46%)	13 (1.93%)	1 (1.61%)	3 (1.90%)	0	0	0	0
Sex								
Male	288 (60.00%)	357 (53.05%)	44 (70.97%)	64 (40.51%)	100 (67.57%)	40 (44.94%)	13 (56.52%)	116 (50.22%)
Female	192 (40.00%)	316 (46.95%)	18 (29.03%)	94 (59.49%)	48 (32.43%)	49 (55.06%)	10 (43.48%)	115 (49.78%)
Ancestry								
African	192 (40.00%)	272 (40.42%)	9 (14.52%)	21 (13.29%)	0	0	0	0
East Asian	10 (2.08%)	16 (2.38%)	6 (9.68%)	16 (10.13%)	0	0	0	0
European	27 (5.62%)	40 (5.94%)	45 (72.58%)	109 (68.99%)	0	0	0	0
Latinx	217 (45.21%)	289 (42.94%)	2 (3.23%)	5 (3.16%)	0	0	0	0
Middle Eastern	30 (6.25%)	52 (7.73%)	0	0	148 (100.00%)	89 (100.00%)	23 (100.00%)	231 (100.00%)
South Asian	3 (0.62%)	4 (0.59%)	0	7 (4.43%)	0	0	0	0
Admixed	1 (0.21%)	0	0	0	0	0	0	0
Comorbidities^a								
Diabetes	194 (40.42%)	214 (31.80%)	25 (40.32%)	45 (28.48%)	75 (50.68%)	11 (12.36%)	9 (39.13%)	unknown
Kidney disease	111 (23.13%)	109 (16.20%)	11 (17.74%)	20 (12.66%)	15 (10.14%)	1 (1.12%)	2 (8.70%)	unknown
Lung disease	139 (28.96%)	142 (21.10%)	14 (22.58%)	25 (15.82%)	18 (12.16%)	3 (3.37%)	2 (8.70%)	unknown
Heart disease	132 (27.50%)	152 (22.59%)	14 (22.58%)	19 (12.03%)	20 (13.51%)	0	1 (4.35%)	unknown
Cancer	131 (27.29%)	174 (25.85%)	5 (8.06%)	7 (4.43%)	5 (3.38%)	0	0	unknown

^aLung disease: asthma, chronic obstructive pulmonary disease, interstitial pulmonary disease, primary pulmonary hypertension; heart disease: coronary artery disease and heart failure.

Table 3. Complete list of all qualifying pLOF variants found in 1,153 COVID-19 cases (673 mild and 480 severe) and 9,589 controls from the Columbia COVID-19 Biobank cohort

Gene	Variant	Effect	HGVS_p ^A	gnomAD exome AF	gnomAD genome AF	Phenotype	Sex	Ancestry cluster
<i>TLR3</i>	4-187004302-C-T	stop_gain	p.Arg488*	1.59×10^{-5}	0	Control	Female	0
<i>TLR3</i>	4-187005327-G-A	splice_donor_variant		4.04×10^{-6}	0	Control	Male	3
<i>TLR3</i>	4-187005911-C-T	stop_gain	p.Arg867*	0	0.000159	Mild COVID-19	Female	1
<i>TLR3</i>	4-187005080-TAGAC-T	frameshift_variant	p.Thr751fs	3.98×10^{-5}	0	Control	Female	7
<i>IRF7</i>	11-613078-G-GA	frameshift_variant	p.Pro439fs	0	3.19×10^{-5}	Control	Male	0
<i>IRF7</i>	11-614300-G-A ^B	stop_gain	p.Gln198*	1.22×10^{-5}	6.38×10^{-5}	Control	Female	0
<i>IRF7</i>	11-614300-G-A ^B	stop_gain	p.Gln198*	1.22×10^{-5}	6.38×10^{-5}	Control	Female	3
<i>IRF7</i>	11-614300-G-A ^B	stop_gain	p.Gln198*	1.22×10^{-5}	6.38×10^{-5}	Control	Female	3
<i>TBK1</i>	12-64879713-CAG-C	frameshift_variant	p.Val421fs	0	3.19×10^{-5}	Control	Male	2
<i>TBK1</i>	12-64882266-G-A	splice_acceptor_variant		0	0	Control	Female	4
<i>IRF3</i>	19-50165422-CCT-C	frameshift_variant	p.Arg255fs	1.19×10^{-5}	0	Control	Male	0
<i>IRF3</i>	19-50166771-CCTGGGG-C	splice_acceptor_variant		0	0	Control	Female	0
<i>IRF3</i>	19-50165291-AGCTCCGCTCACT-A	frameshift_variant	p.Val295fs	3.98×10^{-6}	0	Control	Female	3
<i>IRF3</i>	19-50165422-CCTGT-C	frameshift_variant	p.Asp254fs	1.59×10^{-5}	0	Control	Female	3
<i>IFNAR1</i>	21-34721850-G-A	splice_donor_variant		1.26×10^{-5}	0	Control	Male	1
<i>IFNAR1</i>	21-34713304-G-T	splice_acceptor_variant		0	3.19×10^{-5}	Control	Female	2
<i>IFNAR1</i>	21-34713304-G-T	splice_acceptor_variant		0	3.19×10^{-5}	Control	Female	5
<i>IFNAR1</i>	21-34721439-G-A	stop_gain	p.Trp277*	1.2×10^{-5}	0	Control	Female	5
<i>IFNAR2</i>	21-34619194-CA-C	frameshift_variant	p.Leu128fs	8.03×10^{-6}	3.18×10^{-5}	Control	Female	0
<i>IFNAR2</i>	21-34621013-G-A	splice_acceptor_variant		8.45×10^{-6}	0	Control	Female	0
<i>STAT2</i>	12-56748597-G-A	stop_gain	p.Gln200*	0	0	Control	Male	0
<i>STAT2</i>	12-56748365-G-A	stop_gain	p.Arg223*	0	6.37×10^{-5}	Control	Female	1
<i>STAT2</i>	12-56743896-C-T	stop_gain	p.Trp398*	0	0	Control	Female	2
<i>STAT2</i>	12-56750297-TG-T	frameshift_variant	p.Gln20fs	3.98×10^{-6}	0	Control	Male	2
<i>STAT2</i>	12-56744928-G-A	stop_gain	p.Arg330*	0	0	Severe COVID-19	Female	8

^AHGVS_p uses Ensembl canonical transcripts: *TLR3*-NST00000296795, *IRF7*-ENST00000397566, *TBK1*-ENST00000331710, *IRF3*-ENST00000601291, *IFNAR1*-ENST00000270139, *IFNAR2*-ENST00000342136, and *STAT2*-ENST00000314128. ^BReported in a case by Zhang et al. (1). Ancestry clusters: 0, European; 1, Hispanic/Latinx; 2, African; 3, European; 4, Middle Eastern; 5, Hispanic/Latinx; 6, Middle Eastern; 7, South Asian; 8, East Asian; 9, European.

supplemental material available online with this article; <https://doi.org/10.1172/JCI147834DS1>). In an attempt to replicate the findings by Zhang et al. (1), we performed a gene set-based collapsing-association test stratified by ancestry (2) using 9,589 population controls of similar ancestries (Supplemental Methods). Although exposed asymptomatic individuals would be ideal controls, using population controls will have only a minor impact on power since only a small proportion of exposed individuals develop severe disease. The COVID-19 Host Genetics Initiative has shown that this can be a valid and powerful strategy for host genetic discovery (3).

When testing for pLOF variants with an internal and external minor allele frequency (MAF) of less than 0.001 in any of the 13 genes of interest (Supplemental Methods), we did not detect any enrichment in cases (OR = 1.10 [95% CI 0.03–7.67], $P = 1$). In fact, only one of the cases and 23 controls had a qualifying pLOF variant (0.21% vs. 0.24%). The carrier frequency in controls was dependent on genetic ancestry ranging from 0% to 0.43% across different ancestral clusters. Interestingly, one of the pLOF variants reported by Zhang et al. in a case (1) was present in 3 of our controls (11-614300-G-A in *IRF7*; Table 3). Increasing the MAF threshold to 1% did not increase the number of qualifying pLOF variants. The addition of missense variants also did not improve enrichment (OR = 1.22 [95% CI 0.84–1.71], $P = 0.3$).

We additionally tested 480 patients with severe disease against 673 patients with milder disease that recovered from COVID-19 without the need for intubation, and again observed no clear enrichment (OR = 1.45 [95% CI 0.02–113.51], $P = 1$). Similarly, even when adding missense variants, we did not detect an enrichment (OR = 1.13 [95% CI 0.74–1.72], $P = 0.6$). We also looked for variants in the 13 genes that were listed as pathogenic or likely pathogenic in ClinVar but could not find any in our 480 severe cases. Based on our power calculations, we have over 80% power to detect ORs of 5.5 or greater for rare pLOF variants at an α of 0.05. We are therefore well powered to replicate the findings by Zhang et al. (1) given their reported OR of 8.0 (Figure 1).

The second cohort included patients recruited by the Biobanque Québec COVID-19. In total, 533 participants underwent genome sequencing in partnership with the Canadian HostSeq project. This cohort included 62 severe COVID-19 cases with respiratory failure requiring invasive ventilator support. The control set ($n = 471$) consisted of 158 individuals with mild COVID-19 that did not require ventilator support and 313 participants who were negative for SARS-CoV-2 by PCR. We tested for enrichment of pLOFs in the 62 severe COVID-19 cases compared to the 471 controls (Supplemental Methods) but did not detect a single pLOF in any of the cases or controls, even when we used a more relaxed MAF thresh-

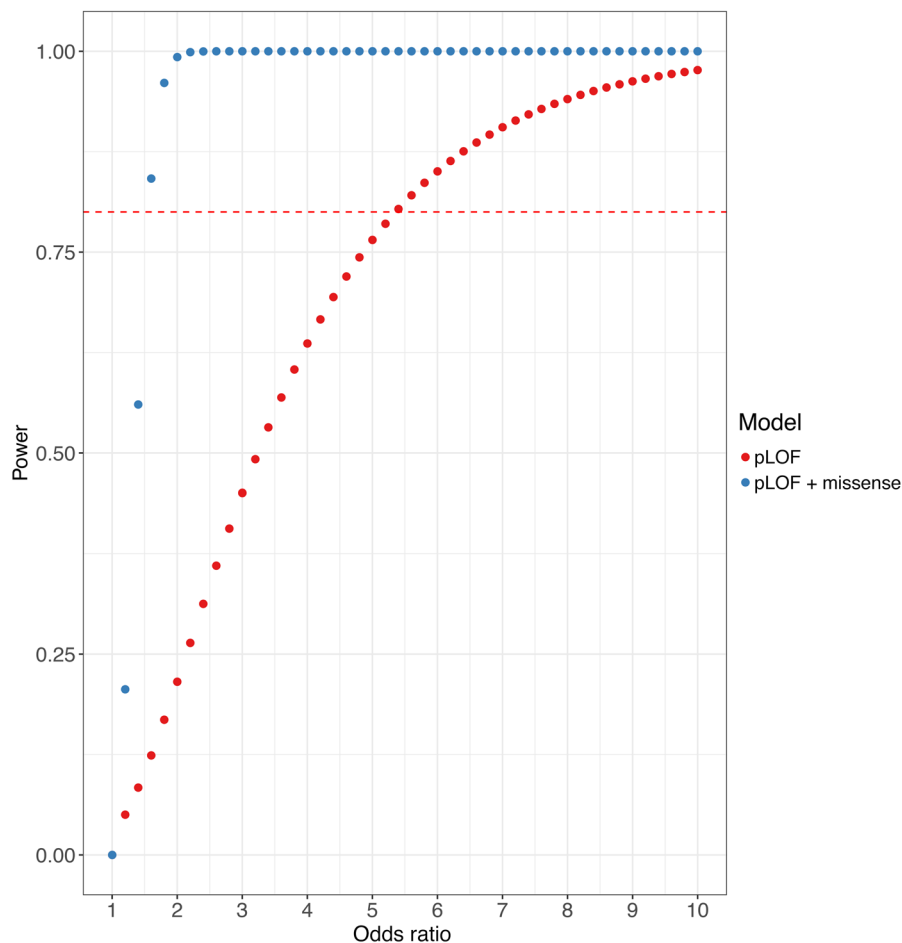


Figure 1. Power curves for a gene set-based collapsing test. Power calculations for the Columbia COVID-19 Biobank cohort of 480 severe cases and 9,589 population controls were performed using the *samplesizeCMH* R package for a dominant model at $\alpha = 0.05$ and a range of odds ratios. Results are shown for the pLOF model and the model including pLOF and missense variants. Because power is influenced by the carrier frequency, we have adequate power to detect effect sizes as small as 1.5 for the model including missense variants compared with 5.5 for the pLOF-only model.

old of 1%. Similarly, we observed no significant enrichment under the missense model (OR = 1.24 [95% CI 0.36–3.46], $P = 0.59$).

The third cohort was recruited in Saudi Arabia and sequenced in partnership with the Saudi Genome project. Exome sequencing was performed in 237 unrelated patients with COVID-19 (148 severe and 89 mild cases; Supplemental Methods). We could not find any cases with a rare pLOF or missense variant in any of the 13 tested genes; even with a MAF of less than 1%, we could only find variants in mildly affected cases.

The fourth cohort was collected by the Qatar Genome Project. Of 14,060 biobank participants with genome sequence data, 700 patients had COVID-19, of which 60 were defined as severe cases and 640 were mild or asymptomatic (see Supplemental Methods). Limiting the analysis to unrelated individuals, there were 5,385 participants including 23 with severe and 231 with mild COVID-19. We did not find any severe cases with rare pLOF or missense variants (MAF < 1%) in the 13 genes analyzed.

In a meta-analysis of all 4 cohorts, we were also not able to detect any significant enrichment for pLOFs (OR = 1.10 [95% CI 0.03–7.67], $P = 1$) or pLOFs and missense variants combined (OR = 1.17 [95% CI 0.83–1.63], $P = 0.34$).

In summary, in our analysis of 4 international COVID-19 biobanks including 1,864 COVID-19 cases (713 severe and 1,151 mild), we could only find a single rare pLOF variant in a severe case and another one in a case with mild disease. We observed no enrichment in pLOF variants in type I IFN immunity genes among

cases of severe COVID-19 compared to mildly affected patients or to ancestry-matched population controls.

These results are further corroborated by a recent analysis of exomes from participants in the UK Biobank performed by Regeneron Inc. (4). The association testing of 1,184 COVID-19 cases versus 422,318 controls showed no association between pLOF variants in these genes after careful adjustment for population stratification and multiple testing. Moreover, there was no association signal when the tests were repeated to include rare deleterious missense variants, or when the case group was limited to severe COVID-19 cases ($n = 471$).

Between this study and the independent report by Regeneron Inc. (4), there were a total of 3,048 cases of COVID-19, including 1,184 cases with severe disease used to test this hypothesis. Despite the large sample size, our collective results do not support the key conclusion by Zhang et al. (1) that 3.5% of severe COVID-19 cases are explained by rare inborn errors in type I IFN immunity genes.

The key differences between our study and the analysis by Zhang et al. (1) are summarized in Supplemental Table 1. Importantly, we have not undertaken any functional studies of missense variants identified in our cohorts. In contrast, Zhang et al. were able to validate 13.7% of missense variants in cases as LOF. Assuming the same rate in our cohorts, however, only 0%–1.5% of our severe cases would harbor an LOF variant, with Europeans showing higher rates compared with Middle Eastern individuals. Our results also demonstrate how important it is to properly con-

trol for ancestry differences between cases and controls, as the frequency of rare variants highly depends on the ancestry of the individuals analyzed. Although the mean age of our case cohort is higher (65.9 years vs. 51.8 years in Zhang et al.), the age of variant carriers reported by Zhang et al. ranged from 17 to 77 years, and 75% of our cases are between 20 and 79. We also recognize that the inheritance pattern for deleterious variants in type I IFN immunity genes is often recessive, and that the penetrance of such variants might be dependent on male sex. However, Zhang et al. found biallelic variants only in a small number of cases and analyzed deleterious variants in both sexes jointly. Taken together, our negative results suggest that the findings by Zhang et al. are not generalizable and highlight the need to rigorously adhere to accepted study design principles when reporting new genetic associations for a set of candidate genes (5, 6).

Methods

Supplemental Figure 1 depicts the full analysis workflow. More details are provided in Supplemental Methods.

Statistics. We used the 2-sided Cochran-Mantel-Haenszel test to test for an association between disease status and the presence of rare variants while controlling for ancestry-cluster membership. For subgroups within a single stratum, we used Fisher's exact test. For all statistical tests, we report uncorrected *P* values, ORs, and 95% CIs. Because the main replication attempt involved a single test, we considered *P* values below 0.05 to be statistically significant.

Study approval. Recruitment and sequencing of participants from the Columbia COVID-19 Biobank were approved by the Columbia University Institutional Review Board (IRB) protocol AAAS7370 and the genetic analyses were approved under protocol AAAS7948. The Biobanque Québec COVID-19 received ethical approval from its multicentric ethics board headed at the Centre Hospitalier de l'Université de Montréal. Recruitment of patients at the Jewish General Hospital (JGH) was also approved by the JGH ethics board. The Saudi Arabia COVID-19 Biobank study was conducted in accordance with the ethical principles of the National Bioethical committee at KACST and approved by the IRB at King Abdullah International Medical Research Centre, Ministry of National Guard - Health Affairs, Riyadh, Ministry of Health, and King Fahad Medical City. The IRBs of all participating hospitals also approved the study. The recruitment and sequencing of participants from the Qatar Biobank (QBB) were approved by the Hamad Medical Corporation Ethics Committee in 2011 and continued with QBB IRB from 2017 onwards and renewed on an annual basis (protocol IRB-A-QBB-2019-0017). The genetic analyses were approved by the QBB IRB protocol E-2020-QBB-Res-ACC-0226-0130.

Author contributions

GP, JBR, AGG, DBG, and KK conceptualized and designed the study. For the Columbia COVID-19 Biobank, NS and KK performed phenotyping; GP and DBG performed sequencing; and GP, CW, AK, DBG, and KK performed genetic analyses. For Biobanque Québec COVID-19, GBL, TN, SZ, VF, VM, and JBR performed phenotyping; RJME, MB, NA, SJMJ, BK, SWS, LJS, PL, JR, GB, VM, and ML performed sequencing; and GBL, TN, SZ, VF, and JBR performed genetic analyses. For the Saudi Genome project, BAS, A Alshareef, MMA, SAA, JA, HMAA, BA, MS Almutairi, ESM, S Massadeh, YMA, N Aljawini, N Albes, M Alaamery, MS

Abedalthagafi, FSAH, LAS, and HEB performed phenotyping; A Almalik, BB, N Aljawini, N Albes, M Alaamery, MS Abedalthagafi, FSAH, EAA, LAS, HEB, JJ, S Massadeh, and MSF performed sequencing; and A Almalik, BB, M Alaamery, MS Abedalthagafi, FSAH, EAA, LAS, S Mangul, and HEB performed genetic analyses. For the Qatar Genome Project, WAM, RB, AAT, and SII performed phenotyping and sequencing, and HM and CS performed genetic analyses. GP performed all cohort meta-analyses. MF performed genetic analyses for the Saudi Genome Project. GP, JBR, AGG, DBG, and KK prepared the manuscript.

Acknowledgments

We would like to thank all study participants for contributing to the biobanks used in this study. Sequencing and phenotyping of the Columbia cohort was made possible by the Columbia COVID-19 Genomics Workgroup, including Andrea Califano, Wendy Chung, Christine K. Garcia, Iuliana Ionita-Laza, Richard Mayeux, Sheila M. O'Byrne, Danielle Pendrick, Muredach P. Reilly, Soumitra Sengupta, Peter Sims, and Anne-Catrin Uhlemann. We thank the Saudi genome members and collaborators including Sameera Al-Johani, Majid Alfadhel, Deema Zahrani, Moneera Alsuwailam, Nouf AlDhawi, Eman Barhoush, Batoul Baz, Faris AlKhlifi, Sarah AlKwai, Maha Albeladi, Faisal Almalki, Iman Mohammad, Ahmed Alaskar, Ebtehal A. Alsolm, Sara S. Alotaibi, Aljohara A. Albabtain, Dona Baraka, Rana Hasanato, Hamza A. Alghamdi, Laila S. Al-Awdah, Amin K. Khattab, Roaa Talal Halawani, Ziab Zakey Alahmadey, Jihad Khalid Albakri, and Walaa Ali Felemban. The authors would like to acknowledge the King Abdullah International Medical Research Center (KAIMRC) for their continuous support. In addition, we thank all the developmental medicine department members for their contributions. We acknowledge King Abdulaziz City for Science and Technology (KACST) for their support through the Saudi Human Genome Project (SHGP) and the Therapy Development Research Program (TDRP). We also thank the Qatar Biobank for data collection and phenotyping. The Qatar Genome Program and Biobank are both Research, Development & Innovation entities within Qatar Foundation for Education, Science and Community Development. We are also thankful to Laith Abu Raddad for EMR data extraction. The Columbia COVID-19 Biobank is supported by the Vagelos College of Physicians & Surgeons Office for Research, Precision Medicine Resource, and Biomedical Informatics Resource of the Columbia University Irving Institute for Clinical and Translational Research (CTSA). Columbia CTSA is funded by the National Center for Advancing Translational Sciences (UL1TR001873). A portion of population controls were from the Washington Heights-Inwood Columbia Aging Project (AG054023). Genome sequencing of Biobanque Québec COVID-19 was funded by the CanCOGeN HostSeq project, with contribution from Fonds de Recherche Québec Santé (FRQS), Génome Québec, and the Public Health Agency of Canada. The Richards group is supported by the Canadian Institutes of Health Research (CIHR), the Lady Davis Institute of the Jewish General Hospital, the Canadian Foundation for Innovation, the NIH, Cancer Research UK, and FRQS. MA and MS Abedalthagafi acknowledge King Abdulaziz City for Science and Technology and the Saudi Human Genome Project for technical and financial support. MMR was supported by Research Support-

ing Project number (RSP-2020/1), King Saud University, Riyadh, Saudi Arabia. GBL is supported by a joint research fellowship from Quebec's Ministry of Health and Social Services, and the FRQS. TN is supported by Research Fellowships of Japan Society for the Promotion of Science (JSPS) for Young Scientists and JSPS Overseas Challenge Program for Young Researchers. JBR is supported by a FRQS Clinical Research Scholarship. This study is part of the www.covid19hg.org consortium.

Address correspondence to: J. Brent Richards, 3755 Cote Ste Catherine, Jewish General Hospital, Montreal, Quebec H3T 1E2, Canada. Phone: 514.340.8222 ext. 24362; Email: brent.richards@mcgill.ca. Or to: David B. Goldstein, 701 West 168th Street, New York, New York 10032, USA. Phone: 212.305.0923; Email: dg2875@cumc.columbia.edu. Or to: Krzysztof Kiryluk, 1150 St. Nicholas Ave, New York, New York 10032, USA. Phone: 212.851.4926; Email: kk473@columbia.edu.

-
1. Zhang Q, et al. Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. *Science*. 2020;370(6515):eabd4570.
 2. Povysil G, et al. Assessing the role of rare genetic variation in patients with heart failure. *JAMA Cardiol*. 2020;6(4):379–386.
 3. The COVID-19 Host Genetics Initiative, Ganna A. Mapping the human genetic architecture of COVID-19 by worldwide meta-analysis [preprint]. <https://doi.org/10.1101/2021.03.10.21252820>. Posted on medRxiv March 12, 2021.
 4. Kosmicki JA, et al. Genetic association analysis of SARS-CoV-2 infection in 455,838 UK Biobank participants [preprint]. <https://doi.org/10.1101/2020.10.28.20221804>. Posted on medRxiv November 3, 2020.
 5. Little J, et al. STrengthening the REporting of Genetic Association studies (STREGA): an extension of the STROBE statement. *Ann Intern Med*. 2009;150(3):206–215.
 6. Povysil G, et al. Rare-variant collapsing analyses for complex traits: guidelines and applications. *Nat Rev Genet*. 2019;20(12):747–759.